# A Generic Polynomial-Time Cell Association Scheme in Ultra-Dense Cellular Networks

Chao Fang[1], Lusheng Wang[1(✉)], Hai Lin[2], and Min Peng[1]

[1] Anhui Province Key Laboratory of Industry Safety and Emergency Technology, School of Computer Science and Information Engineering, Hefei University of Technology, Hefei, China
wanglusheng@hfut.edu.cn

[2] Key Laboratory of Aerospace Information Security and Trusted Computing, Ministry of Education, School of Cyber Science and Engineering, Wuhan University, Wuhan, China

**Abstract.** Cell association in heterogeneous cellular networks is a significant research issue, but existing schemes mainly optimize a single objective and could not solve such a problem with a generic utility function in polynomial time. This paper proposes a cell association scheme for generic optimization objectives with polynomial-time complexity, which employs a virtual base station method to transform it into a 2-dimensional assignment problem solved by Hungarian algorithm. Based on this scheme, a framework for the tradeoff among multiple optimization objectives is designed. This framework jointly considers spectral efficiency and load balancing, designs a weight factor to adjust their impacts on the optimization, and uses an experience pool to store the relationship between performance demands and corresponding weight factor values. For an instantaneous cell association decision in a given network scenario, the association results are obtained as soon as the corresponding factor value is taken from the pool and the Hungarian algorithm is called for the matching. Compared with existing schemes, our proposal achieves a better tradeoff between system capacity and UE fairness with an extremely low time cost.

**Keywords:** Heterogeneous cellular networks · Cell association · 2-dimensional assignment problem · Hungarian algorithm · Fairness

## 1 Introduction

To solve problems caused by the increment of traffic load and the lack of wireless resource, cellular networks evolve toward heterogeneity integrating femtocells with traditional macrocells, called heterogeneous cellular networks (HCNs) [1]. The deployment of femtocells brings in an augmentation of the system capacity thanks to the small-scale reuse of resource [2], but the association problem


H. Gao et al. (Eds.): ChinaCom 2019, LNICST 312, pp. 417–432, 2020.
https://doi.org/10.1007/978-3-030-41114-5_31

between all the base stations (BSs) and user equipments (UEs) becomes more and more complex when the densification of BSs increases. Therefore, it becomes a critical issue to find a fast cell association scheme in HCNs, which should achieve a good tradeoff between multiple performance metrics.

In the literature, there are many studies on cell association in HCNs. Some of them just considered a traditional performance metric, such as signal-to-interference-plus-noise ratio (SINR), so that an association strategy with high system capacity could be quickly obtained [3]. However, due to the difference of transmission powers between a femtocell and a macrocell, too many UEs tended to access the macro one, making the traditional SINR-based scheme unsuitable for HCNs. To find a suitable solution for HCNs, cell association and scheduling were jointly optimized in [4], which transformed it into a distributed convex optimization and used an alternating direction method of multipliers to solve it. This algorithm associated more UEs to underloaded femtocells to improve load balancing and the throughput on cell edges. [5] proposed a cell association scheme based on UE behavior awareness, which obtained the association result based on UEs' instantaneous states (such as their deployment and mobility features) and cells' characteristics, so that the scheme could dynamically approach the network's maximum throughput.

There are some other studies on the optimization methods of cell association. [6] considered the problem with the proportional-fair utility function and transformed it into a convex optimization by relaxing the binary variables representing the associations into continuous variables between 0 and 1, which was then solved by the Lagrange duality method. [7] jointly optimized cell association and power control to maximize the system total utility and minimize the power consumption. The problem was modeled as a mixed integer convex optimization by an annealing-based coalition game and the primal decomposition theory. [8] proposed a cell association scheme based on online Q-learning, which continuously learned UE behaviors and the dynamic UE environment, so that load balancing was improved under the premise of guaranteeing UE quality of service (QoS). [9] proposed a deep Q-learning based scheme, which achieved an optimal association under the premise of guaranteeing the downlink UE QoS.

None of the above schemes is polynomial, but some existing studies transform cell association into an assignment problem that is solvable in polynomial time. [10] considered a virtual base station (VBS) idea to transform the problem into the association between UEs and VBSs, where one BS was mapped into a number of VBSs, hence becoming a 2-dimensional assignment problem. UEs wanted to maximize their own profits and BSs wanted load balancing, so a Nash bargaining game was used to model the conflicts between UEs and BSs. [11] jointly optimized the cell association problem with BS dormancy and considered a utility function obeying proportional fairness. On the one hand, the cell association subproblem was transformed into a 2-dimensional assignment problem and solved by Hungarian algorithm. On the other hand, a low-complex algorithm based on a successive approximation method was proposed for joint optimization. [12] jointly considered cell association and almost blank subframe (ABS) ratio as a combinatorial optimization problem. For a given ABS ratio, Hungarian algorithm was used to match UEs and VBSs, and finally a strategy corresponding to a relatively small ABS ratio but a large number of associated UEs was obtained.

In summary, existing cell association schemes mainly optimize a single performance metric and most of them are not polynomial. Some polynomial-time schemes only work for proportional-fair utility functions. Therefore, this paper proposes a generic polynomial-time scheme and uses it as the core of a cell association framework that optimizes multiple performance metrics with a tradeoff. In details, a VBS method is employed to transform cell association into a 2-dimensional assignment problem between UEs and VBSs. Then, a weighting factor is used to adjust the impacts of spectral efficiency and load balancing. We store the relationship between achieved performance and corresponding factor values in an experience pool. Once an association decision is required, an association result is obtained by running Hungarian algorithm on a virtual weight matrix that is calculated based on the corresponding factor taken from the experience pool. The advantages of the proposal are threefold: its complexity is polynomial, its objective function could be generic, and it achieves a better tradeoff among multiple performance metrics.

The remainder of this paper is organized as follows. Section 2 provides the system model. Section 3 describes the proposed scheme and the framework. Section 4 shows the simulation results. In the end, the paper is concluded in Sect. 5.

## 2 System Model

We consider a circular region covered by a macrocell and a number of femtocells. The BS of the former is in the center and the BSs of the latter are deployed in the circle obeying a certain distribution, as shown in Fig. 1. $\mathbf{BS} = \{BS_j | j = 1, \ldots, N\}$ is used to denote all of them, where $N$ is the total number of BSs in the whole region. UEs in the area are represented by $\mathbf{UE} = \{UE_i | i = 1, \ldots, M\}$, where $M$ is the total number.

The scheme in this paper is designed to be generic for a series of utility functions and for different levels of fairness, as explained at the end of Sect. 3.1, but to simplify the description of the proposal and to make it easy to follow, we model the problem here as the maximization of the system capacity, given by

$$\max_{x_{ij}} \sum_{j \in \mathbf{BS}} \sum_{i \in \mathbf{UE}} x_{ij} C_{ij} \tag{1}$$

$$s.t. \sum_{j \in \mathbf{BS}} x_{ij} = 1 \ \forall i \in \mathbf{UE} \tag{1a}$$

$$x_{ij} = \{0, 1\} \ \forall i \in \mathbf{UE}, j \in \mathbf{BS} \tag{1b}$$

where $C_{ij} = B \cdot s_{ij} / \sum_{i \in \mathbf{UE}} x_{ij}$ is the capacity of the link between $UE_i$ and $BS_j$, $B$ is the total bandwidth that a BS possesses in the system. $x_{ij}$ represents the association between $UE_i$ and $BS_j$, i.e., $x_{ij} = 1$ if it is associated and 0 otherwise,

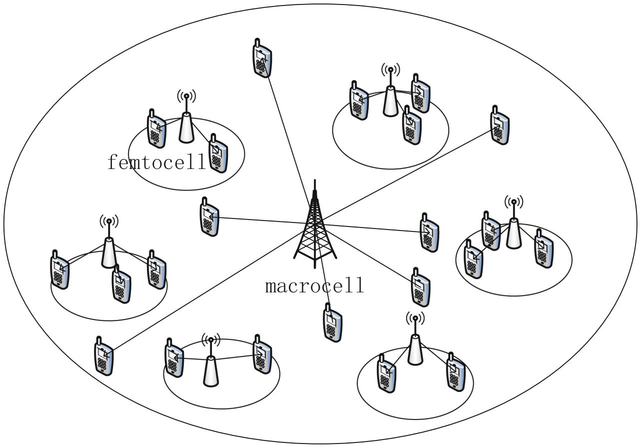



**Fig. 1.** System model.

so $\sum_{i \in \mathbf{UE}} x_{ij}$ represents the total number of UEs associating with $BS_j$. (1b) indicates that one UE can only associate with a single BS. $s_{ij}$ represents the spectral efficiency of $UE_i$ associating with $BS_j$ which can be written as

$$s_{ij} = \log_2 \left( 1 + \frac{Pr_{ij}}{I_j + N_0} \right) \tag{2}$$

where $N_0$ is the variance of the additive white Gaussian noise (AWGN). $Pr_{ij}$ is the reception power for the link between $UE_i$ and $BS_j$, and it is calculated by

$$Pr_{ij} = Pt_j - PL_{ij} \tag{3}$$

where $Pt_j$ is the transmission power for the link between $UE_i$ and $BS_j$, and $PL_{ij}$ is the link's pathloss. $I_j$ represents the total interference from the other cells

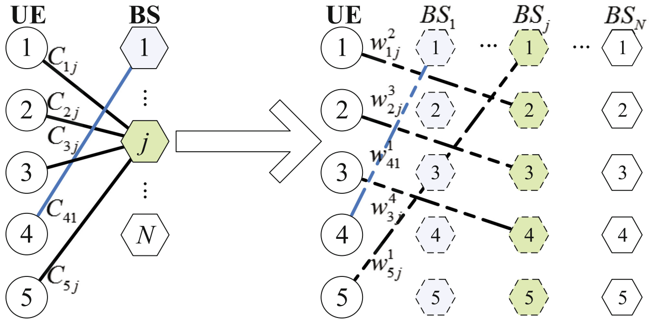
$$I_j = \sum_{\beta \in \mathbf{BS}/j} Ir_j(\beta) \tag{4}$$

where $Ir_j(\beta)$ represents the interference from cell $\beta$ to $BS_j$. Note that, the calculation of $Ir_j(\beta)$ is different for uplink and downlink, so the values in the formed weighting matrix are different, as well as the association results, but it does not affect much of the proposal in this paper. Meanwhile, we do not combine uplink and downlink for an integrated decision, because traffic loads of uplink and downlink may go through different BSs in HCNs.

## 3 Proposed Scheme and Framework

### 3.1 Proposed Generic Polynomial-Time Scheme

The system capacity maximization problem modeled by (1) is a typical one-to-multiple assignment problem. To the best of our knowledge, it cannot be solved in polynomial time if the utility function is not in a proportional-fair form [10,11]. We employ the VBS concept in [11] and try to find a heuristic method to transform the problem with generic utility functions to a 2-dimensional assignment problem, so that it can be solved in polynomial time by a traditional method, called Hungarian algorithm.




(a) Associations between UEs and BSs (b) Associations between UEs and VBSs

**Fig. 2.** An example of VBS method.

The VBS method maps each BS into $M$ VBSs and each VBS can only be associated with one single UE. In this way, we may transform the one-to-multiple assignment problem between UEs and BSs into a 2-dimensional assignment problem between UEs and VBSs, as shown in Fig. 2. Lines in Fig. 2a represent the associations between UEs and BSs, and symbols marked on the lines denote these links' capacities. Similarly, spotted lines in Fig. 2b represent the associations between UEs and VBSs, and symbols marked on the spotted lines denote the utilities of these associations, i.e., $w_{ij}^l$ denotes the utility obtained by associating $UE_i$ with the $l$th VBS of $BS_j$.

To make sure the objective function of the transformed problem is equivalent to the original in (1), the total capacities of the UEs associated with each BS should equal to the summation of the utilities of these UEs associated with VBSs (Condition 1). Taking Fig. 2 as an example, we should have $C_{1j} + C_{2j} + C_{3j} + C_{5j} = w_{1j}^2 + w_{2j}^3 + w_{3j}^4 + w_{5j}^1$ and $C_{41} = w_{41}^1$. In the meantime, since the method used to solve the transformed association problem is the Hungarian algorithm which chooses in priority the VSB providing a larger utility for each UE, $w_{ij}^l$ should be monotonously decreasing with the increasing of $l$ (Condition 2). Finally, $w_{ij}^l$ is used to calculate the input matrix of the Hungarian algorithm, so each $w_{ij}^l$ for any $i$, $j$, and $l$ should be known before an association result is obtained. In other words, $w_{ij}^l$ should not be related to other UEs' features, such as their capacities and their associated VBSs, because you do not know if they are associated with the same BS during the Hungarian algorithm. Therefore, $w_{ij}^l$ should be an expression only related to $UE_i$, $BS_j$, and the VBS index $l$ (Condition 3).

To design an expression of $w_{ij}^l$ that fits for all the above conditions simultaneously is a mission impossible, otherwise problem (1) should have been already precisely transformed into a 2-dimensional assignment problem and solved. Inspired by the design of $w_{ij}^l$ for the problem with a proportional-fair objective function in [11], we relax Condition 1 by giving the summation of $w_{ij}^l$ a range that may contain the total capacity of these UEs, so a heuristic design is obtained as

$$w_{ij}^l = \begin{cases} s_{ij} & l = 1 \\ s_{ij} + k \times [(l-1)\log(l-1) - l\log(l)] & 2 \le l \le M \end{cases} \tag{5}$$

where $k \in [0, +\infty)$ is a weighting factor between spectral efficiency and fairness. When $k = 0$, $w_{ij}^l$ in (5) is decided by $s_{ij}$. Thus, UEs all choose the BSs providing them the highest spectral efficiencies, leading to load imbalance and poor UE fairness. When $k$ increases, the importance of the second part in (5) increases and the impact from $s_{ij}$ gradually decreases, making the difference between associating with different BSs smaller. When $k$ becomes quite large, $w_{ij}^l$ is mainly decided by the second part in (5). Even though the UEs may be distributed asymmetrically, they tend to be averagely assigned to the BSs, making UE fairness varies without caring about their locations. In a word, $k = 0$ and $k = +\infty$ are two extremes representing the considerations of only spectral efficiency and of only averaging the number of UEs among the BSs, so there must be a $k$ that achieves a good tradeoff between the two objectives. Based on the VBS method and the designed $w_{ij}^l$ in (5), problem (1) is transformed into a 2-dimensional assignment problem as follows:

$$\max_{x_{ij}^l} \sum_{j \in \mathbf{BS}} \sum_{i \in \mathbf{UE}} \sum_{l=1}^{M} x_{ij}^l w_{ij}^l \tag{6}$$

$$s.t. \sum_{j \in \mathbf{BS}} \sum_{l=1}^{M} x_{ij}^l = 1 \; \forall i \in \mathbf{UE} \tag{6a}$$

$$\sum_{i \in \mathbf{UE}} x_{ij}^l \le 1 \; \forall j \in \mathbf{BS}, 1 \le l \le M \tag{6b}$$

$$x_{ij}^l \in \{0, 1\} \; \forall i \in \mathbf{UE}, \forall j \in \mathbf{BS}, 1 \le l \le M \tag{6c}$$

where $x_{ij}^l$ represents the association between $UE_i$ and the $l$th VBS of $BS_j$, i.e., $x_{ij}^l = 1$ if it is associated and 0 otherwise. (6a) guarantees that each UE only associates with one single VBS, (6b) guarantees that each VBS is only associated with one UE. In the meantime, note that (1) is also a utility function integrating the two objectives, so there is probably a $k$ making (6) almost equivalent to (1).

Based on the above analysis and design, problem (1) is heuristically transformed into a 2-dimensional assignment problem. Then, Hungarian algorithm is employed to solve it with the following key steps:

(a) for each BS, the utility between each of its VBS and each UE is calculated, and an $M$-by-$M$ utility matrix is obtained. For the $N$ BSs, we obtain $N$ $M$-by-$M$ utility matrices and joint them as an $M$-by-$MN$ long square matrix $\mathbf{W}$. Denoting its entry on the $i$th row and the $y$th column as $W(i, y)$, it can be represented by $W(i, y) = w_{ij}^l$, where $j = y$ ceil $M$, $l = y$ mod $M$.
(b) we use the Hungarian algorithm on $\mathbf{W}$ to obtain the associations between UEs and VBSs, which can be simply transformed back to the associations between UEs and BSs.

Note that, although (1) is modeled as a capacity maximization problem, the proposed scheme is generic for utility functions with a form of $\sum_{i \in \mathbf{UE}} x_{ij} f(\cdot)$, as well as different fairness levels integrated on it. The function is a summation of the utilities of all the UEs, $\{x_{ij} | i \in \mathbf{UE}, j \in \mathbf{BS}\}$ are the binary variables representing the associations, and $f(\cdot)$ should be a utility function integrating the concept that the resource of a BS is averagely divided by all its associated UEs. All kinds of utility functions obeying these conditions can be solved by our proposal. Meanwhile, the effect of the weighting factor $k \in [0, +\infty)$ integrated in our scheme similarly corresponds to the effect of $\alpha \in [0, +\infty)$ in the well-known $\alpha$-fairness concept, so our scheme could equivalently solve the utility functions integrating different levels of fairness by taking a corresponding $k$ value. In summary, our scheme is generic for all kinds of utility functions obeying the above conditions and different levels of fairness.

### 3.2 Usage of the Proposed Scheme in a Two-Stage Framework

The above subsection is described in a theoretical way, but the usage of such a method for cell association in a real network should be explained and one key problem must be further considered, i.e., the process to obtain a suitable $k$ is too slow for an instantaneous cell association decision. Therefore, a cell association framework is proposed in this subsection, which uses the above theoretical method as the core and achieves a tradeoff among multiple performance metrics. This framework divides the cell association process into two stages: one preprocessed experimental stage, as shown by the left part of Fig. 3, to obtain an experience pool containing a number of representative $k$ values and corresponding performance metric values, and one trigger-based instantaneous decision stage, as shown by the right part of Fig. 3, to quickly reach an appropriate association result.

The experimental stage may be run in a centralized manner by a macro BS or in a distributed manner by a number of cloud computers. In details, it first deploys BSs and UEs to form a similar network as the real scenario, and each BS is mapped to $M$ VBSs. For each $k$ value, it then obtains the long square utility matrix $\mathbf{W}$ by (5) and calls Hungarian algorithm to reach an association result. Finally, the performance in terms of various metrics for this association is calculated. After a large number of simulation rounds are completed, the average performance metric values for each $k$ are calculated and stored into the experience pool which is actually a table containing the $k$ values and their corresponding performance metric values. Since this pool is decided by

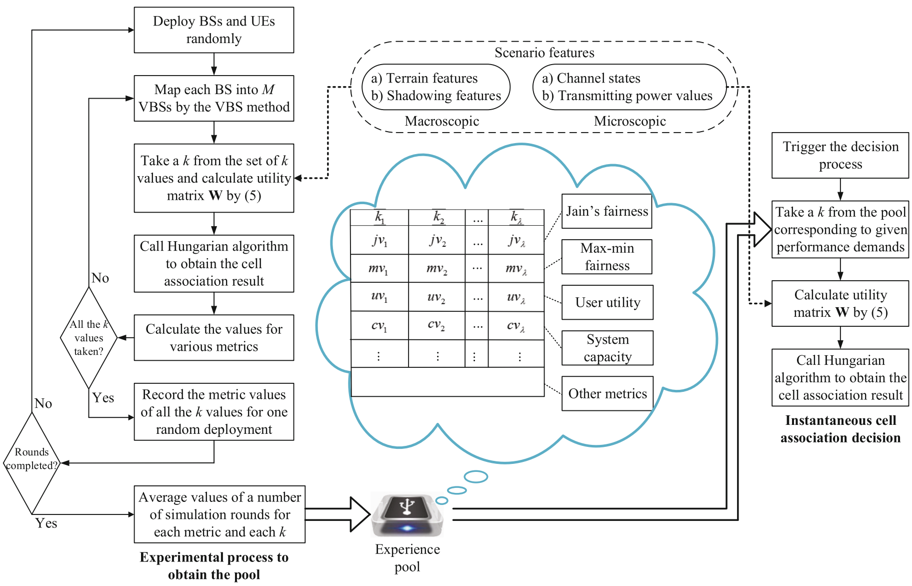



**Fig. 3.** A two-stage framework using the proposed cell association scheme.

the macroscopic features of the scenario, it is updated only when the scenario encounters an obvious change.

After the pool is obtained, it is ready for usage during an instantaneous cell association decision. The decision stage may be triggered periodically or by some obvious changes, such as the movement of some representative UEs, but the triggering issue is out of the scope of this study. Once the decision process is triggered, a $k$ value corresponding to the demanded performance is taken from the experience pool. Then, (5) is used to calculate the long square utility matrix $\mathbf{W}$ and Hungarian algorithm is used to reach an association result for this instantaneous decision. Note that, the receiving power values of useful signals and interferences for the calculation of $\mathbf{W}$ in the experimental process is different from those values in the decision process. For the former, the useful signals and interferences are calculated based on the channel model and the randomly deployed BSs and UEs in the simulation. For the latter, they are calculated based on the real values evaluated by channel estimation in the network, i.e., the microscopic features of the scenario shown by Fig. 3.

## 4 Performance Evaluation

In our simulations, BSs and UEs are distributed within a circular region with a radius of 25 m. For the sake of limited space, only downlink channel features are used for the calculation of utility matrices, and the channel is modeled by the close-in free space reference distance model with frequency-dependent path loss exponent for 5G scenarios [13]:

$$PL(f,d)[\text{dBm}] = 20\log_{10}\left(\frac{4\pi f}{c}\right) + 10n\left[1 + b\left(\frac{f-f_0}{f_0}\right)\right]\log_{10}\left(\frac{d}{1m}\right) + X_\sigma \tag{7}$$

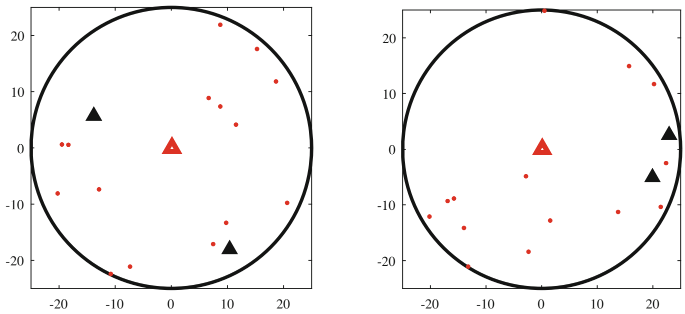
where $f$ is the carrier frequency, $n$ is the path loss exponent, $b$ is a slope parameter, $X_\sigma$ represents the shadowing, and $f_0$ is the reference frequency. Detailed parameter values are listed in Table 1.

**Table 1.** Simulation parameters.

| Parameter | Value |
|---|---|
| Circular region radius | 25 m |
| Femtocell transmission power | 21 dBm |
| Macrocell transmission power | 30 dBm |
| Femtocell bandwidth | 6 MHz |
| Macrocell bandwidth | 20 MHz |
| Variance of AWGN | −174 dBm/Hz |
| Carrier frequency | 3.5 GHz |
| Path loss exponent | 2.59 |
| Slope parameter | 0.01 |
| Shadowing | 7.4 dBm |
| Reference frequency | 39.5 GHz |

### 4.1 Simulations of the Preprocessed Experimental Stage

In this subsection, the preprocessed experimental stage in the framework is simulated. According to our experience, $k$ in (5) should take representative values to form an appropriate experience pool, but here we take values $\{0, 0.05, 0.1, 0.3, 0.5, 0.7, 0.9, 1, 1.2, 1.4, 1.6, 1.8, 2, 4, 6, 8, 10, 20, 30, 35, 40\}$ so that the changes of the curves are clearly demonstrated. For each $k$, performances should be averaged by a large number of simulation rounds, such as 500 in our simulation. We simulate two scenarios, one heterogeneous network scenario and one homogeneous network scenario. At the beginning of 5G network construction, femtocells cannot be densely deployed and macrocells should be still used as a main bearer for traffic loads, forming an uncrowded heterogeneous network. Therefore, the simulated heterogeneous network is composed of 1 macro BS in the center (the big red triangle), 2 femtocells (the small black triangles), and 15 uniformly-distributed UEs. The 2 femtocells' locations may be random and changeable during the 500 simulation rounds, as shown in Fig. 4a or fixed on the right corner as shown in Fig. 4b.




(a) Femtocell random deployment. (b) Femtocell special deployment.

**Fig. 4.** BS deployments in the heterogeneous network scenario.

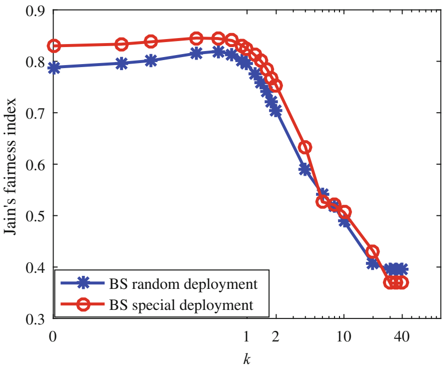



**Fig. 5.** UE fairness for heterogeneous network scenario.

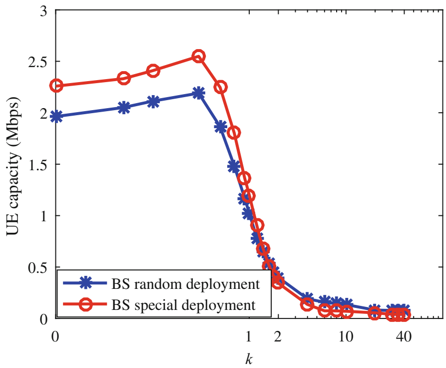



**Fig. 6.** Minimum UE capacity for heterogeneous network scenario.

Figure 5 shows the Jain's fairness values of the two types of BS deployments. Along with the increase of $k$, the trends of UE fairness curves generally increase at first and then decrease. As explained in Subsect. 3.1, it is not difficult to understand that $k = 0$ and $k = +\infty$ are two extremes resulting in low fairness, so there must be a $k$ in the middle corresponding to the maximum Jain's fairness. Seen from Fig. 5, the $k$ should be both around [0.3, 0.5] for the two types of deployments of the simulated heterogeneous network scenario. Figure 6 shows the minimum UE capacity of all the UEs. We find that, the trends of the curves are similar to those of fairness, and the best $k$ for this performance metric should be also around 0.3.

The system utilities obtained by (5) are shown in Fig. 7. Since the part multiplied on $k$ is negative in (5), the curves always decrease with the increase of $k$. Figure 8 shows the system capacities of the two types of deployments. We can see that, the curves gradually increase with the increase of $k$. Since bandwidth

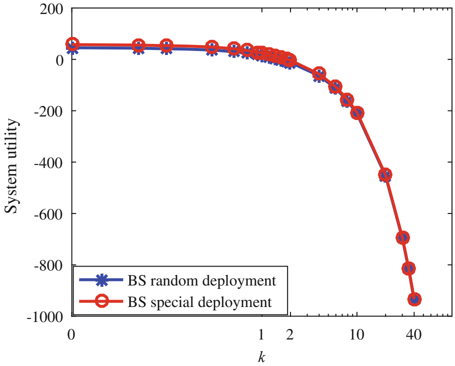



**Fig. 7.** System utility for heterogeneous network scenario.

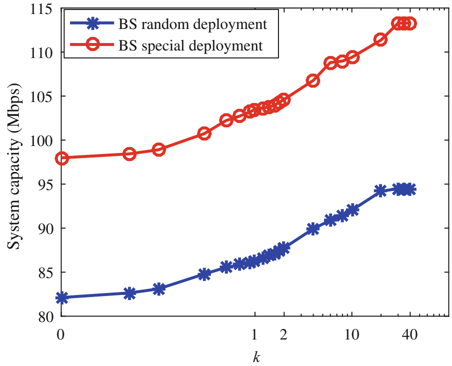
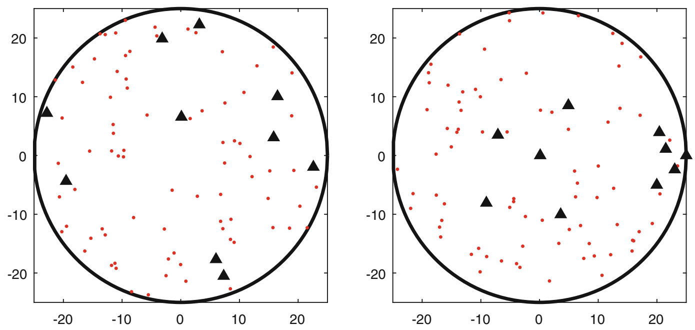



**Fig. 8.** System capacity for heterogeneous network scenario.

and transmission power of the macrocell are both much larger than those of femtocells, most of the UEs tend to choose the macrocell when $k = 0$, which can be easily understood by checking (5). This is actually inbeneficial to the system capacity due to the fact that too many UEs share the bandwidth of the macrocell. When $k$ increases, some UEs gradually change to choose the femtocells, improving the system capacity until the association becomes average among the three BSs.

The simulated homogeneous network scenario is shown by Fig. 9, where 10 femtocells and 80 UEs are deployed in the circular region, and two types of deployments are considered, i.e., random deployment and special deployment with 5 femtocells on the right corner. This scenario may represent the case when 5G network is fully constructed, so femtocells are dense enough to afford all the traffic loads and macrocells are free for network management only.



(a) Femtocell random deployment. (b) Femtocell special deployment.

**Fig. 9.** BS deployments in the homogeneous network scenario.

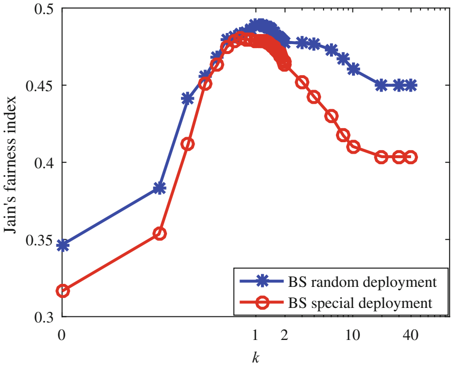



**Fig. 10.** UE fairness for homogeneous network scenario.

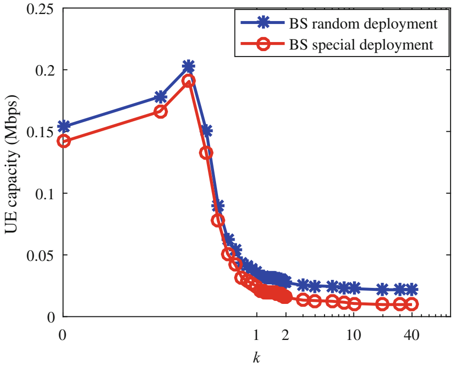



**Fig. 11.** Minimum UE capacity for homogeneous network scenario.

Generally speaking, UE fairness, minimum UE capacity, system utility, and capacity of the homogeneous scenario have similar trends as the heterogeneous one, as shown in Figs. 10, 11, 12, and 13, but there are also some differences. One is the Jain's fairness values when $k$ is relatively small. For the heterogeneous scenario, most UEs tend to choose the macro BS when $k$ is small as explained above. Meanwhile, we note that the simulation area is within a 25-meter circle, so the UEs are all relatively close to the macro BS with small path loss values. Now that these UEs share the bandwidth of the macro BS averagely and their path loss values are all small, they tend to obtain similar capacities, leading to large Jain's fairness. By contrast, UEs in the homogeneous scenario tend to choose different femtocells and the transmission power of femto BSs is relatively small, so they tend to obtain obviously different capacities, leading to small Jain's fairness, as shown in Fig. 10. Also note that for different scenarios, the $k$ values corresponding to the maximum Jain's fairness could be different, and for the homogeneous scenario it should be around $k = 1$ as shown in Fig. 10.

The one of the two deployments that leads to a better performance is also quite different for the two scenarios. For the heterogeneous scenario, special deployment obviously achieves a better performance in terms of most evaluated metrics, while the homogeneous scenario is generally inverse. Based on our massive experiments, we find that the main reason for this phenomenon is still the bandwidth and the transmission power of the macro BS. Since the macro BS is quite aggressive for attracting UEs to associate, it seems beneficial to the whole system to put the two femto BSs far from it in the small simulation region, such as on the edge. By contrast, the femto BSs in the homogeneous scenario are with the same bandwidth and transmission power, so the results for this scenario does not show this feature.

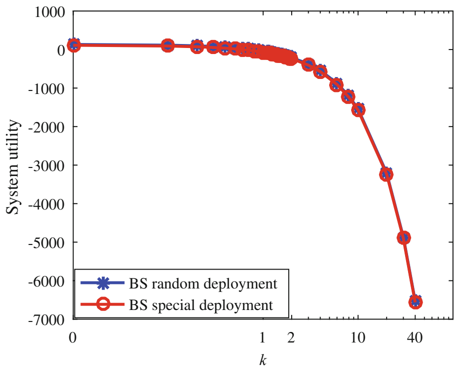



**Fig. 12.** System utility for homogeneous network scenario.

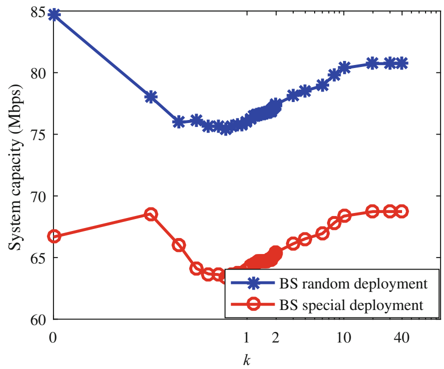



**Fig. 13.** System capacity for homogeneous network scenario.

### 4.2 Simulations of Cell Association Decision and Comparisons with Other Schemes

This subsection simulates the cell association decision stage in the framework and evaluates the performance of the final association results. A $k$ value should be taken from the experience pool, so that multiple performance metrics could be comprehensively considered by a tradeoff or a quite high performance for a certain metric is reached as an objective performance for comparison. To compare with our scheme, we select the max SINR scheme in [3], the simulated annealing based scheme in [7], and the Q-learning based scheme in [8] for the following simulations. Besides, the heterogeneous network scenario with randomly deployed femtocells in Subsect. 4.1 is selected due to the fact that existing related works mainly consider heterogeneous networks. Based on the simulation results in Subsect. 4.1, we select $k = 0.3$ in our scheme for the comparisons with the other schemes. This $k$ value emphasizes UE fairness and the capacity of the worst UE due to the fact that compared schemes generally consider system capacity more than fairness.

Figure 14 shows the Jain's fairness values of various schemes. The proposed scheme achieves a high UE fairness. $k = 0.3$ in the simulation corresponds to a very high UE fairness already demonstrated by Fig. 5, and here we also find that the proposed scheme can get higher UE fairness when optimizing UE fairness alone in this figure. Figure 15 shows the minimum UE capacities (max-min fairness) of various schemes. Similar to UE fairness, our scheme also achieves a very good result. Note that, our scheme is better than Q-learning scheme and Max-SINR scheme in terms of UE fairness and minimum UE capacity.

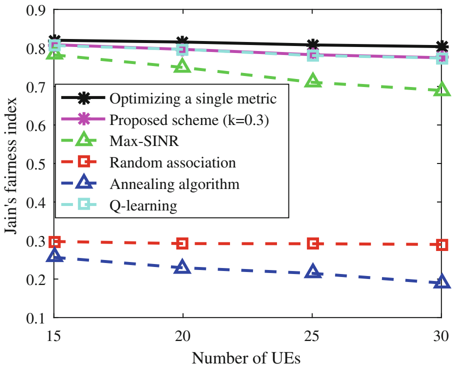



**Fig. 14.** UE fairness values of various schemes.

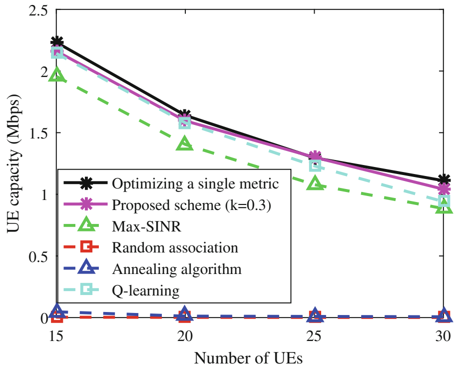



**Fig. 15.** Minimum UE capacities of various schemes.

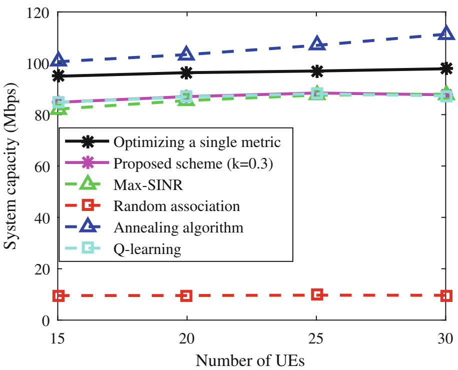



**Fig. 16.** System capacities of various schemes.

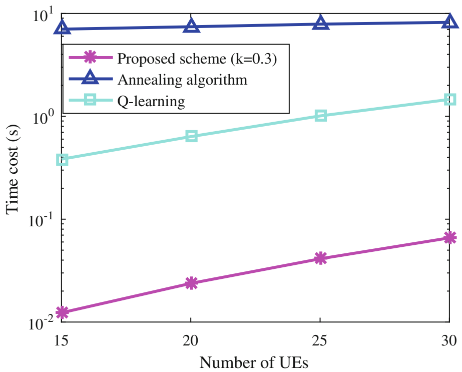



**Fig. 17.** Time costs of various schemes.

Figure 16 shows the system capacities of various schemes. Simulated annealing takes a quite long time to search for a near-optimal solution, so its achieved system capacity is obviously better than the others. Q-learning based scheme and our proposed scheme both achieve relatively high system capacity. Note that, we set $k = 0.3$ that is a value corresponding to high fairness, but the achieved system capacity is still quite competitive in the simulated schemes. If a different $k$ is set for optimizing system capacity only, our scheme achieves an obviously better system capacity, as shown by the curve "Optimizing a single metric".

Figure 17 shows the time costs of three schemes. Simulated annealing based scheme is undoubtedly slow, and the proposed scheme is obviously faster than the Q-learning based scheme. The reason is that, Q-learning requires too many iterations before finding a good solution, but our scheme requires only solving a 2-dimensional assignment problem with polynomial-time complexity.

## 5 Conclusion

This paper studied the cell association problem in heterogeneous cellular networks and a generic polynomial-time scheme was proposed. On the one hand, the scheme employed the VBS method to transform heuristically the problem into a form that was solvable in polynomial time. On the other hand, the scheme achieved a tradeoff among multiple performance metrics by a two-stage framework. An experience pool containing a series of $k$ values and corresponding performance metric values was used to link the two stages. Simulation results showed that the proposed scheme achieved a better tradeoff between spectral efficiency and UE fairness with an extremely low time cost. For the sake of limited space, this paper only simulated the case using downlink channel features for the calculation of the utility matrix, and simulations of the case using uplink channel features will be a future work.

**Acknowledgements.** This work was funded by the Fundamental Research Funds for the Central Universities of China under grant no. PA2019GDQT0012.

## References

1. Liu, D., et al.: User association in 5G networks: a survey and an outlook. IEEE Commun. Surv. Tutor. **18**(2), 1018–1044 (2016)
2. Andrews, J., Claussen, H., Dohler, M., Rangan, S., Reed, M.: Femtocell: past, present, and future. IEEE J. Sel. Areas Commun. **30**(3), 497–508 (2012)
3. Andrews, J., Singh, S., Ye, Q., Lin, X., Dhillon, H.: An overview of load balancing in HetNets: old myths and open problems. IEEE Wirel. Commun. **21**(2), 18–25 (2014)
4. Ge, X., Li, X., Jin, H., Cheng, J., Leung, V.: Joint user association and user scheduling for load balancing in heterogeneous networks. IEEE Trans. Wirel. Commun. **17**(5), 3211–3225 (2018)
5. Sun, Y., Feng, G., Qin, S., Sun, S.: Cell association with user behavior awareness in heterogeneous cellular networks. IEEE Trans. Veh. Technol. **67**(5), 4589–4601 (2018)
6. Shen, K., Yu, W.: Distributed pricing-based user association for downlink heterogeneous cellular networks. IEEE J. Sel. Areas Commun. **32**(6), 1100–1113 (2014)
7. Qian, L., Wu, Y., Zhou, H., Shen, X.: Joint uplink base station association and power control for small-cell networks with non-orthogonal multiple access. IEEE Trans. Wirel. Commun. **16**(9), 5567–5582 (2017)
8. Li, Z., Wang, C., Jiang, C.: User association for load balancing in vehicular networks: an online reinforcement learning approach. IEEE Trans. Intell. Transp. Syst. **18**(8), 2217–2228 (2017)
9. Zhao, N., Liang, Y., Niyato, D., Pei, Y., Wu, M., Jiang, Y.: Deep reinforcement learning for user association and resource allocation in heterogeneous cellular networks. IEEE Trans. Wirel. Commun. (in press)
10. Wang, W., Wu, X., Xie, L., Lu, S.: Femto-matching: efficient traffic offloading in heterogeneous cellular networks. In: IEEE INFOCOM, pp. 325–333. IEEE, Hong Kong (2015)

11. Prasad, N., Arslan, M., Rangarajan, S.: Exploiting cell dormancy and load balancing in LTE HetNets: optimizing the proportional fairness utility. IEEE Trans. Commun. **62**(10), 3706–3722 (2014)
12. Mishra, S., Rangineni, S., Murthy, C.: Exploiting an optimal user association strategy for interference management in HetNets. IEEE Commun. Lett. **18**(10), 1799–1802 (2014)
13. 5GCM. http://www.5gworkshops.com/5GCM.html